APPLICATION

FOR

UNITED STATES LETTERS PATENT

SPECIFICATION

TO ALL WHOM IT MAY CONCERN:

Be it known that, Dongshan Fu, Gopinath Kuduvalli, and Shehrzad Qureshi, citizens of The People's Republic of China, Canada, and the United States of America, respectively, and residing in Santa Clara, CA, San Jose, CA, and Palo Alto, CA, respectively, have invented certain improvements in a **APPARATUS AND METHOD FOR REGISTERING 2D RADIOGRAPHIC IMAGES WITH IMAGES RECONSTRUCTED FROM 3D SCAN DATA** of which the following description in connection with the accompanying drawings is a specification, like reference characters on the drawings indicating like parts in the several figures.

1

APPARATUS AND METHOD FOR REGISTERING 2D RADIOGRAPHIC IMAGES WITH IMAGES RECONSTRUCTED FROM 3D SCAN DATA

FIELD OF THE INVENTION

[0001] The present invention relates to a method and system for registering 2D (two-dimensional) radiographic images with images reconstructed from 3D (three-dimensional) scan data. More particularly, the invention relates to a method and system for registering 2D stereo x-ray images data with digitally reconstructed radiographs of 3D CT scan data.

BACKGROUND

[0002] Medical image registration is useful in many areas of medicine, for example radiosurgery. In radiosurgery, tumors and other lesions are treated by delivering a prescribed high dose of high-energy radiation to the target area, while minimizing radiation exposure to the surrounding tissue. Radiosurgery therefore calls for an ability to accurately focus on a target region, so that only the target receives the desired high doses of radiation, while surrounding critical structures are avoided. Typically, 3D imaging modalities, such as computed tomography (CT), magnetic resonance (MR) imaging, or positron emission therapy (PET) are used to generate diagnostic 3D images of the anatomical region containing the targeted area, for treatment planning purposes. These tools enable practitioners to identify the anatomical organs of a patient, and to precisely locate any abnormalities such as tumors.

[0003] To correct patient position or align radiation beam, the change in target position at the time of treatment (as compared to the position at the time of the diagnostic treatment planning) needs to be detected. This is accomplished by registering the 2D image acquired at the treatment time with the 3D CT scan obtained at the time of treatment planning.

[0004] The target positions are defined using the 3D diagnostic CT scan by physicians at the time of treatment planning. CT scans allow an image of the internal structure of a target object to be generated, one cross-sectional slice at a time. The CT data is used as the reference to determine the patient position change during treatment. Typically, synthesized 2D images such as digitally reconstructed radiographs (DRRs) are generated from the 3D CT data, and are used

as 2D reference images. Similarity measures are used to compare the image intensity in the x-ray and the DRR images, in order to determine the patient pose change. In the field of medical image registration, this problem is categorized as a 2D/3D registration.

[0005] The methods used in the 2D/3D registration procedure can be divided into two categories. The first category includes methods based on image features. The image features may be anatomical edges, or segmented objects. Registration accuracy depends on the accuracy of edge detection, or the accuracy of object segmentation. The main advantage of feature-based methods is computation speed. Because the full information content of the image is not used, however, accuracy is sacrificed. The second category includes methods based on image intensity content. In intensity-based methods, the original images are used for the registration process. Therefore, a good accuracy can usually be achieved. Because a lengthy computation time is required, however, intensity-based methods are not practical for purposes of radiosurgery, or for clinical practice in general.

[0006] Image-guided radiosurgery requires precise and fast positioning of the target at the treatment time. In practice, the accuracy should be below 1 mm, and the computation time should be on the order of a few seconds. Unfortunately, it is difficult to meet both requirements simultaneously, because of several reasons. First, the two different modality images, i.e. CT scan images and x-ray images, have different spatial resolution and image quality. Generally, x-ray image resolution and quality are superior to the resolution and quality of DRR images. Second, DRR generation relies on a proper attenuation model. Because attenuation is proportional to the mass intensity of the target volume through which the beam passes through, the exact relationship between the traversed mass intensity and the CT image intensity needs to be known, in order to obtain an accurate modeling. Establishing this relationship is difficult, so the linear attenuation model is often used. However, the skeletal structures in DRR images cannot be reconstructed very well using the linear model, the DRRs being only synthetic x-ray projection images. Finally, x-ray images usually have a large image size (512 x 512). For better registration accuracy, it is desirable to use the full resolution image. Full resolution images are rarely used, however, due to the extremely slow computation that results from using such

images.

[0007] U.S. Pat. No. 5,901,199 by Murphy et al. (the "Murphy patent") describes a high-speed inter-modality image registration via iterative feature matching. The Murphy patent is a feature-based method. Prior to treatment, extraction and segmentation of silhouettes of the patient's skull are performed in order to make a feature mask. A set of DRR images are generated from the 3D CT data and are then masked, in order to isolate key pixels that are associated with anatomical edge features. The masked image contains only 5%-10% of the total image pixels. During treatment, the acquired x-ray images are similarly masked. The registration is conducted on the masked DRRs and the masked X-ray images. The registration process is completed in a few seconds. However, the accuracy and stability of the estimates are not sufficient to meet the sub-mm precision that is required in radiosurgery applications.

[0008] For these reasons, there is a need for a method and system for performing 2D/3D medical image registration using as little computing time as possible, while at the same time meeting the requisite accuracy for radiosurgical applications.

SUMMARY OF THE INVENTION

[0009] The present invention is directed to a precise and rapid method and system for registering 2D x-ray images with images that have been reconstructed from 3D scan data. A hierarchial and iterative framework for the registration algorithm permits a higher accuracy to be achieved, in both the translational and rotational adjustments. The total computation time is about an order of magnitude faster than other techniques existing in the prior art.

[0010] A method is provided for registering one or more 2D x-ray images of a treatment target with one or more 2D reconstructed images of the target that are based on previously generated 3D scan data. Preferably, the 2D reconstructed images are DRRs, generated using the same positions and angles of the imaging beams used for the x-ray images. The method includes performing in-plane rotations of the DRRs within the image plane of the x-ray images, thereby generating reference DRRs. The x-ray images are processed so that the orientation, image size, and bit depth of the x-ray images match the orientation, image size, and bit depth of the reference DRRs.

[0011] The method involves determining the value of parameters (x, y, θ) and (r, φ) that are required for registering the x-ray image of the target with the reference DRRs of the target. (x, y, θ) represent the in-plane translational and rotational parameters within the image plane of the x-ray images, (x, y) indicating the requisite amount of translation within the image plane in the directions of the x- and y- axes, respectively, and θ indicating the requisite amount of rotation within the image plane. (r, φ) represent the out-of-plane rotational parameters, and indicate the requisite amount of out-of-plane rotations about mutually orthogonal axes that are defined in a 3D coordinate system, and that are orthogonal to the image plane.

[0012] In order to determine these parameters, a 3D multi-level matching is first performed, in order to determine an initial estimate for the in-plane transformation parameters (x, y, θ). Based on these parameters (x, y, θ) obtained by 3D multi-level matching, an initial 1-D search is performed for each of the pair of out-of-plane rotation parameters (r, φ). The in-plane translation parameters (x, y) are then refined, using 2D sub-pixel matching, to increase the accuracy of these

parameters.

[0013] The in-plane rotation parameter (θ) is then refined, based on the out-of-plane rotation parameters (r, φ) obtained from the initial 1D search, and on the updated in-plane transformation parameters (x,y), in order to increase the accuracy of the in-plane rotation parameter φ. 1D interpolation is used in this step.

[0014] Next, each of the out-of-plane rotation parameters (r, φ) are refined separately, based on the refined in-plane translation and rotation parameters. The refining steps are iteratively repeated, until a predetermined accuracy is reached. Finally, the out-of-plane rotation parameters (r, φ) are refined, using 1D interpolation, in order to achieve the desired resolution.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] FIG. 1 illustrates the in-plane translational and rotational parameters (x, y, θ), and the out-of-plane rotational parameters (r, φ), for registering a 2D radiographic image with previously generated 3D scan data.

[0016] FIG. 2 illustrates the generation of 2D DRRs from 3D CT scan data of a treatment target within an anatomical region of a patient.

[0017] FIG. 3 illustrates a flowchart of the registration algorithm used in a 2D/3D registration method, in accordance with one embodiment of the present invention.

[0018] FIG. 4 illustrates a multi-resolution image representation for a multi-level matching process used to estimate the in-plane transformation parameters.

[0019] FIG. 5 illustrates a schematic block diagram of an apparatus for performing 2D/3D registration, in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION

[0020] The present invention is directed to an improved method and system for performing medical 2D/3D registration. The tracking method and system of the present invention is useful in radiosurgery and radiotherapy; however the method and system of the present invention can also be used in applications other than radiosurgery and radiotherapy, i.e. in any application where there a need to track rigid object by registering 2D radiographic images onto 3D scan data. While 2D x-ray images are described in the preferred embodiment of the present invention, the present invention can be used for the registration of any other type of 2D images onto 3D scan data.

[0021] In overview, the method of the present invention includes generating a reconstructed 2D image from the 3D scan data. The change in the position of the target (or other rigid object) in the radiographic image, as compared to the position of the target in the 3D scan data (as indicated in the reconstructed 2D image) is described using 3D rigid body transformations. The 2D x-ray image is registered onto the reconstructed 2D image, by determining the value of 3D rigid body transformation parameters that represent the difference in the position of the target as shown in the x-ray image as compared to the position of the target as shown by said 2D images reconstructed from said 3D scan data.

[0022] In radiosurgery, 2D/3D registration is necessary, for example, in order to correct patient position or properly align the radiosurgical beam. It is accomplished by registering the X-ray image acquired at the treatment time with the 3D CT scan obtained at the treatment planning time. In radiosurgery, the target positions are defined using a diagnostic 3D scan by physicians at the time of treatment planning. A CT scan is most frequently used, and will be described in the exemplary embodiments discussed in this section; however, it should be noted that other 3D scanning methods, such as MRI, ultrasound, or PET scanning, may also be used in other embodiments of the present invention. The 3D data is used as reference, to determine the patient position change during treatment. For this purpose, 2D reference images are reconstructed from the 3D scan data. Preferably, digitally reconstructed radiographs (DRR) are generated from 3D

8

CT data, and are used as 2D reference images.

[0023] Because there are no external forces imposed on a treatment target during radiation treatment, it is reasonable to treat the treatment target as a rigid body, i.e. an object whose internal geometric relationships remain static or unchanged over time. The 3D rigid transformation can be described using the six degrees of freedom: three translations of the CT center and three rotations (roll, pitch and yaw) about three orthogonal axes. In the present invention, two orthogonal x-ray projections are utilized, in order to determine the six degrees of freedom, i.e. to solve for these six parameters (x, y, z, r, p, w). The registration in each projection is performed individually, and the results of the registration for each projection are subsequently combined, to obtain the six 3D transformation parameters.

[0024] The 2D x-ray projection images of the target are formed by transmitting imaging beams (having a known intensity, and having known positions and angles with respect to the target), generated from a respective pair of x-ray sources, through the target and onto cameras A and B. Each x-ray image A and B are characterized by a respective image plane, defined by mutually orthogonal x- and y- axes in a coordinate frame defined by the two x-ray cameras A and B: x_A and y_A for projection A, and x_B and y_B for projection B.

[0025] FIG. 1 illustrates the geometry of the in-plane translational and rotational parameters (x, y, θ), and the out-of-plane rotational parameters (r, φ), in an embodiment of the invention in which a pair of orthogonal 2D projection images, A and B, are used. As illustrated in FIG. 1, a 3D CT coordinate system, i.e. a coordinate system for the target as viewed in the frame of the CT scan study (taken at the time of treatment planning), can be defined. During treatment, the patient assumes a position within the real-time camera coordinate frames (defined by the two x-ray cameras A and B, respectively), that does not necessarily match the position of the patient as seen from within the 3D CT coordinate system. The differences in the position and orientation of the anatomical target images within the respective radiographs (real-time x-ray, versus DRR) correspond to the difference in the three-dimensional position and orientation of the target between the camera- and the CT coordinate frames, and are found by solving for the parameters

(x, y, z, r, p, w).

[0026] In the embodiment illustrated in FIG. 1, the x-axis in the 3D CT coordinate system is directed inward into the paper, and is not referenced. The 2D projections A and B are viewed from the directions o_As_A and o_Bs_B respectively. The directions of axis x_A in the 2D projection in the camera A coordinates, and axis x in the 3D CT coordinates are opposite to each other. The direction of axis x_B in the camera B coordinates, and axis x in the 3D CT coordinates, are the same.

[0027] As shown in FIG. 1, each projection is characterized by a respective set of transformation parameters, namely $(x_A, y_A, \theta_A, r_A, \varphi_A)$ for projection A, and $(x_B, y_B, \theta_B, r_B, \varphi_B)$ for projection B. The two out-of-plane rotations (with respect to the image plane) in projections A and B are denoted by (r_A, φ_A) and (r_B, φ_B) respectively, where r denotes the amount of rotation about the x-axis (in the 3D CT coordinate system), and φ denotes the amount of rotation about the o_As_A axis (for projection B) or the o_Bs_B axis (for projection A). The in-plane translations and rotation in projections A and B are denoted $(x_A y_A \theta_A)$ and $(x_B y_B \theta_B)$, respectively. As easily seen from FIG. 1, $x_A y_A$ and $x_B y_B$ denote the amount of translations within the image planes for each projection (A and B) in the directions of the x- and y- axes that define each image plane (x_A- and y_A- for projection A, and x_B- and y_B- for projection B), while θ_A and θ_B denote the amount of rotation within each image plane about an axis (not shown) that is perpendicular to both the x_A- (or x_B-) and y_A- (or y_B-) axes.

[0028] As can be seen from FIG. 1, the out-of-plane rotation φ_A in projection A is the same as the in-plane rotation θ_B in projection B, and the out-of-plane rotation φ_B in projection B is the same as the in-plane rotation θ_A in projection A. The use of the two projections A and B thus over-constrains the problem of solving for the six degrees of freedom. It can be seen from FIG. 1 that $x_A = x_B$, $r_A = r_B$, $\theta_A = \varphi_B$ and $\theta_B = \varphi_A$.

[0029] For projection A, given a set of reference DRR images which correspond to different combinations of the two out-of-plane rotations (r_A, φ_A), the 2D in-plane transformation $(x_A, y_A,$

φ_A) can be estimated by the 2D image comparison. Determining the two out-of-plane rotations (r, θ_B) relies on which reference DRR is used for best similarity match. Similarly, the 2D in-plane transformation (x_B, y_B, θ_B) and the out-of-plane rotations (r_B, φ_B) can be estimated for projection B.

[0030] In one embodiment of the present invention, digitally reconstructed radiographs (DRRs) are generated offline from the 3D CT scan, and used as references for determining the position of the patient and the target. In an exemplary embodiment, a set of reference DRRs is calculated that samples the full range of possible patient orientations. For example, the reference DRRs are calculated before treatment, by varying the orientation of the CT volumetric image of the target relative to the camera projection geometry. The reference DRRs are then stored for use in a lookup table during treatment. A comparison is made between the radiographs, acquired in real time, with each of the DRRs, to select the DRR that best matches the relevant real-time radiograph.

[0031] FIG. 2 illustrates the generation of a 2D DRR from 3D CT scan data of a treatment target within an anatomical region of a patient. In FIG. 2, the volumetric 3D CT image of the target is referred to with the aid of reference numeral 60. The DRRs 65A and 65B, shown in FIG. 2, are artificial, synthesized 2D images that represent the radiographic image of the target that would be obtained, if imaging beams were used having the same intensity, position and angle as the beams used to generate the real time x-ray projection images, and if the target were positioned in accordance with the 3D CT scan data. In other words, the DRRs are calculated from prior 3D CT data, in an exact emulation of the real-time camera perspectives. The reference numerals 50A and 50B illustrate the hypothetical positions and angles from which the imaging beams would be directed through a target positioned in accordance with the CT volumetric image 60 of the target.

[0032] Typically, DRRs are generated by casting hypothetical beams or rays through the CT volumetric image of the target. Each ray goes through a number of voxels of the 3D CT image 60. By integrating the CT numbers for these voxels along each ray, and projecting onto an

imaging plane (shown as 70A and 70B, respectively, in FIG. 2), the resultant image would emulate the radiograph that would be obtained by passing rays from hypothetical camera locations and angles (shown schematically as 50A and 50B, respectively) through a target positioned in accordance with the volumetric 3D image 60. Ray tracing algorithms, known in the art, are generally used to generate the DRRs.

[0033] Fig. 3 illustrates a flowchart of the registration algorithm used in a 2D/3D registration method, performed in accordance with one embodiment of the present invention. In the present invention, the registration algorithm is designed in a hierarchical and iterative framework. The registration algorithm permits an accurate and rapid registration of a 2D x-ray image of a target with previously generated 3D scan data of the target.

[0034] As seen from FIG. 3, the first step (step 110 in FIG. 3) in the registration process is to pre-process the raw x-ray images, so that the orientation, image size, and bit depth of the x-ray image emulates the orientation, image size, and bit depth of the reconstructed 2D image.

[0035] In the embodiment of the present invention illustrated in FIG. 3, the registration process is described in terms of six distinct phases (illustrated in FIG. 3 as steps 120, 130, 140, 150, 160, and 170). In phase 1 (step 120 in FIG. 3), the in-plane transformation parameters (x, y, θ) are initially estimated using a set of in-plane rotated DRR images, which are generated offline from the nominal reference DRR (in 0 degree). The most intensive computation in the registration process is the computation of the in-plane rotation. To achieve a rapid computation, it is desirable to compute as many in-plane rotations as possible for the reference DRRs, before starting the registration process. The process of generating in-plane rotated DRRs is thus carried out offline, after the reference DRRs for out-of-plane rotations are generated. All the reference DRR images are stored in memory, and used for registering each real-time x-ray image that is acquired during patient alignment and treatment.

[0036] In step 120, the three parameters are rapidly searched using a 3D multi-level matching method (described in connection with FIG. 4 below). A sum of absolute differences method ("SAD"; described in the co-pending ACCL-127 application) is used as the similarity measure.

In this step, there is no floating computation. The pixel accuracy for the translations (x,y) and half-degree accuracy for the in-plane rotation (θ) are achieved.

[0037] In the next step, i.e. step 130 (phase 2 of the registration process), the two out-of-plane rotations (r, φ) are separately searched in one dimension, based on the values of the in-plane parameters (x, y, θ), determined in previous step 120. A more complicated similarity measure, based on pattern intensity (described in the co-pending ACCL-127 application), is used to detect the reference DRR image that corresponds to a combination of two out-of-plane rotations (r, φ). The search space for the possible rotation angles is the full search range of out-of-plane rotation angles. For an initial estimate, the full search range is sampled at every one-degree interval. In step 140 (phase 3), the in-plane translation parameters (x, y) are refined using 2D sub-pixel matching. 2D sub-pixel matching is a full range search method. Based on the updated transformation parameters (x, y, θ, r, φ) obtained from the previous step in the registration, a set of DRR images (3 x 3 or 5 x 5) is generated by translating the unknown reference DRR, one sub-pixel at a time. The in-plane translations (x, y) in sub-pixel accuracy are refined by finding the best match between the x-ray image and the DRR images.

[0038] In step 150 (phase 4), the in-plane rotation parameter θ is refined using 1D interpolation, based on the updated values for the in-plane translation parameters (x, y) from step 140, and the updated values of the out-of-plane rotation parameters (r, φ) from step 130. In step 160 (phase 5), the out-of-plane rotations are separately refined to a better accuracy using 1D search, based on the updated values for the in-plane transformation parameters (x, y, θ), from steps 140 and 150. In steps 140, 150, and 160 (phases 3, 4, and 5), a similarity measure method based on pattern intensity (described in the co-pending ACCL-127 application) is used, to ensure higher accuracy.

[0039] Steps 140, 150, and 160 are iteratively repeated until, a sufficient accuracy is obtained. Once the desired accuracy is reached, the final out-of-plane rotations are 1D interpolated, in the final step 170 (6th and last phase) of the registration process.

[0040] FIG. 4 illustrates a multi-resolution image representation for the multi-level matching

13

process, used in the first phase (step 120 in FIG. 3) to initially estimate the in-plane transformation parameters. The full-size image is at the bottom (Level 1). The upper images (Level 2, Level 3 and Level 4) have lower spatial resolution. The lower resolution images are obtained by lower pass filtering, and by sub-sampling of the full-size images.

[0041] As a fast search method, multi-level matching is used for an initial estimate the in-plane transformation parameters. The basic idea of multi-level matching is to match the images at each level successively, starting with the lowest image resolution level (Level 4). The results at the lower resolution level serve to provide rough estimates for the in-plane transformation parameters (x, y, θ). The output at a lower level is then passed to the subsequent level characterized by a higher resolution. The parameters (x, y, θ) are refined, using the higher resolution images. In the final results obtained through multi-level matching, the accuracy of the translations depends on the spatial resolution of the image having the highest resolution (Level 1). The accuracy of the rotations depends on the sampling intervals of the in-plane rotations, during the DRR initialization process described in paragraph 35 above.

[0042] There may be some risks inherent in multi-level matching. The estimates at lower levels may fall within local minima, and far away from global minima. In this case, further matching at subsequent levels (at higher resolutions) may not converge to the global minima. To overcome this risk, multiple candidates of estimates are used. Many candidates for an optimal matching at a lower level are passed on to the higher resolution level. The higher the number of candidates used, the more reliable are the estimates. The best candidates are ranked by the SAD values.

[0043] In FIG. 4, denoting the full image size in Level 1 by $W \times H$, the image sizes are $\frac{W}{2} x \frac{H}{2}, \frac{W}{4} x \frac{H}{4}$ and $\frac{W}{8} x \frac{H}{8}$ in Level 2, Level 3 and Level 4, respectively. For translations, the search range in the lowest resolution level is the full search range that is calculated from the difference between the DRR and x-ray image sizes. Because of the smallest image size $\frac{W}{8} x \frac{H}{8}$ at the lowest level, the full range search can be completed in a very short time. The same small search range is (-2, +2) pixels for the remaining resolution levels. Because of the small search

range, the search can be completed quickly, even at large image sizes. For the rotations, the search range in the lowest resolution level is a full search range, at a denser sampling rate. In the higher resolution levels, partial search ranges are used, at a less dense sampling rate.

[0044] FIG. 5 illustrates a schematic block diagram of an apparatus 200 for performing 2D/3D registration, in accordance with one embodiment of the present invention. In overview, the apparatus 200 includes a means 201 for generating pre-treatment 3D scan data of the target; a radiation source 202; an imaging system 206; and a controller 208. The means 201 is a 3D scanner (for example a CT scanner, or an MRI scanner, or a PET scanner). The radiation source generates at least one radiographic imaging beam having a known intensity, and having a known position and angle relative to the target.

[0045] The controller 208 includes means for generating a set of 2D DRR images of the target, using the 3D scan data from the CT scanner, and using the known location, angle, and intensity of the imaging beam generated by the radiation source. The controller 208 also includes software for determining a set of in-plane transformation parameters (x, y, θ) and out-of-plane rotational parameters (r, φ), the parameters representing the difference in the position of the target as shown in the x-ray image, as compared to the position of the target as shown by the 2D reconstructed images.

[0046] The controller 208 further includes 1) software for performing a 3D multi-level matching to determine an estimate for the in-plane transformation parameters (x, y, θ); 2) software for performing a 1-D search for each of the pair of out-of-plane rotation parameters (r, φ), based on the estimated in-plane parameters (x, y, θ); and 3) software for iteratively refining the in-plane parameters (x, y, θ) and the out-of-plane parameters (r, φ), until a desired accuracy is reached.

[0047] In practice, a high accuracy is obtained for both translations and rotations after just a few iterations, using the method and system of the present invention. For translations, an accuracy of 0.5mm or better is reached, and for rotations, an accuracy of 0.5 degrees or better is reached. The total computing time is a few seconds, which is an order of magnitude faster than other methods in the prior art.

[0048] While the invention has been particularly shown and described with reference to specific preferred embodiments, it should be understood by those skilled in the art that various changes in form and detail may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.